

UNITED STATES PATENT APPLICATION

for

METHOD AND APPARATUS FOR A MULTI-USER VIDEO NAVIGATION
SYSTEM

Inventors:

Bo Shen

Susie Wee

Prepared by:

WAGNER, MURABITO & HAO LLP

TWO NORTH MARKET STREET

THIRD FLOOR

SAN JOSE, CALIFORNIA 95113

(408) 938-9060

METHOD AND APPARATUS FOR A MULTI-USER VIDEO NAVIGATION
SYSTEM

TECHNICAL FIELD

5 The presently claimed invention relates to the field of
video transmission systems. More specifically, the
presently claimed invention relates to a multi-user video
navigation environment.

10 BACKGROUND ART

Video transmission systems present a video stream of
data. Each video stream contains a sequence of individual
picture frames that each contain a still image. When the
picture frames are shown together sequentially, a video
15 sequence of picture frames is presented. Transmission of a
single video sequence can occur over a communication
network to a single client device or multiple client
devices.

20 Heretofore, functional control over transmission of
the video sequence was only possible through a single
controlling device. For example, a conventional approach
can provide multiple users with navigational functionality.
However, only a single user is allowed to have functional
25 control over transmission of the video sequence. In one
case, the controlling user can specify the position of the
source camera to determine content of the video sequence.
The captured video sequence is then compressed with

compression hardware or software and transmitted to the multiple client devices displaying the video sequence.

In addition, further processing of the video sequence, e.g., resizing of the transmitted video sequence, is again controlled by a single device. As such, multiple client devices receiving transmission of the video sequence all display the same video sequence.

The aforementioned approach to controlling functionality over transmission of the video sequence is not scalable to multiple users. Only one user can control the content of the transmitted video sequence. The remaining client devices receiving transmission of the video sequence are essentially dummy devices that have no functional control over the transmission of the video sequence. As such, all the client devices are limited to the view determined by the position of the camera, which can only be controlled by one client device at a time.

Thus, a need exists for overcoming singular functional control of video presentation for a video sequence. Another need exists for scalable functional control of video presentation for a video sequence.

DISCLOSURE OF THE INVENTION

The present invention provides a method and system for multi-user video navigation. One embodiment of the present invention provides a method that achieves the above
5 accomplishment and which also provides for overcoming the singular functional control of video presentation for a video sequence. Additionally, one embodiment of the present invention provides a method that achieves the above accomplishments and which also provides for scalable
10 functional control of video presentation over a video sequence.

Specifically, one embodiment of the present invention discloses a method of video presentation whereby multiple
15 users can each view and navigate cropped windows of a live-captured or stored video sequence in compressed form. A video sequence is captured from a stationary camera. For live captured video, an uncompressed version and compressed version following a compression format are captured. For
20 stored video, a decoder or digitizer module replaces the capture module to obtain the pixel domain information. In response to user requests, cropped windows can be navigated from one location to another within the dimensions of the video sequence. For each cropped window, information data
25 from the compressed and uncompressed versions are combined to generate a windowed portion of the video sequence that is in the same compression format. Whenever possible, data from the compressed version is used, and data from the uncompressed version is substituted when necessary to

generate each windowed portion of video sequence.
Transmission of each windowed portion of the video sequence
to the various requesting client devices occurs
contemporaneously.

5

These and other technical advantages of the present
invention will no doubt become obvious to those of ordinary
skill in the art after having read the following detailed
description of the preferred embodiments which are
illustrated in the various drawing figures.

10

0597098-10008148

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a block diagram of an exemplary electronic device capable of providing multiple and navigable windows in a video sequence, in accordance with one embodiment of the present invention.

Figure 2 illustrates a diagram of an exemplary communication system capable of capturing a video sequence and providing multiple and navigable windows of the video sequence to a plurality of client devices, in accordance with one embodiment of the present invention.

Figure 3 illustrates a block diagram of an exemplary frame in a video sequence showing a cropped window, in accordance with one embodiment of the present invention.

Figure 4 is a diagram showing an exemplary adjusted motion vector for a block within a frame in a video sequence that has moved from a first location to a second location, in accordance with one embodiment of the present invention.

Figure 5 is a flow diagram illustrating steps in a method for generating a compressed video stream for a portion of a video sequence that corresponds to a stationary window in said video sequence, in accordance with one embodiment of the present invention.

Figure 6 is a flow diagram illustrating steps in a method for generating a compressed video stream for a

portion of a video sequence that corresponds to a window that is navigating from a first location to a second location in said video sequence, in accordance with one embodiment of the present invention.

5

Figure 7 is a flow diagram illustrating steps in a method for generating a window in a video sequence and navigating the window to a second location in the video sequence, in accordance with one embodiment of the present invention.

10

HP-10008148/JPW/LCH

BEST MODES FOR CARRYING OUT THE INVENTION

Reference will now be made in detail to the preferred embodiments of the present invention, a method for providing multiple and navigable portions of a video sequence, examples of which are illustrated in the accompanying drawings. While the invention will be described in conjunction with the preferred embodiments, it will be understood that they are not intended to limit the invention to these embodiments. On the contrary, the invention is intended to cover alternatives, modifications and equivalents, which may be included within the spirit and scope of the invention as defined by the appended claims.

Furthermore, in the following detailed description of the present invention, numerous specific details are set forth in order to provide a thorough understanding of the present invention. However, it will be recognized by one of ordinary skill in the art that the present invention may be practiced without these specific details. In other instances, well known methods, procedures, components, and circuits have not been described in detail as not to unnecessarily obscure aspects of the present invention.

NOTATION AND NOMENCLATURE

Some portions of the detailed descriptions which follow are presented in terms of procedures, steps, logic blocks, processing, and other symbolic representations of operations on data bits that can be performed on computer memory. These descriptions and representations are the means used by

those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. A procedure, computer executed step, logic block, process, etc., is here, and generally, conceived to be a self-consistent sequence of steps or instructions leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated in a computer system. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussions, it is appreciated that throughout the present invention, discussions utilizing terms such as "forming," or "generating," or "cropping," or "navigating," or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers

or other such information storage, transmission or display devices.

COMPUTER SYSTEM ENVIRONMENT OF THE PRESENT INVENTION

5 Referring now to Figure 1, portions of the present invention are comprised of computer-readable and computer-executable instructions which reside, for example, in computer-readable media of an electronic system 100 that is capable of generating multiple and movable portions of a
10 video sequence. Figure 1 is a block diagram of interior components of an exemplary electronic system 100, upon which embodiments of the present invention may be implemented.

Figure 1 illustrates circuitry of an exemplary
15 electronic system 100. Exemplary electronic system 100 includes an address/data bus 120 for communicating information, a central processor 101 coupled with the bus 120 for processing information and instructions, a volatile memory 102 (e.g., random access memory (RAM), static RAM
20 dynamic RAM, etc.) coupled with the bus 120 for storing information and instructions for the central processor 101, and a non-volatile memory 103 (e.g., read only memory (ROM), programmable ROM, flash memory, EPROM, EEPROM, etc.) coupled to the bus 120 for storing static information and
25 instructions for the processor 101.

Exemplary electronic system 100 also includes a data storage device 104 (e.g., memory card, hard drive, etc.) coupled with the bus 120 for storing information and

instructions. Data storage device 104 can be removable. Exemplary electronic system 100 also contains an optional electronic display device 105 coupled to the bus 120 for displaying information to a user. The display device 105
5 utilized with the electronic system 100 may be a liquid crystal device, cathode ray tube (CRT), field emission device (FED, also called flat panel CRT) or other display device.

10 With reference still to Figure 1, a signal Input/Output device 108 which is coupled to bus 120 for providing a communication link between electronic system 100 and a wired or wireless network environment is described. As such, signal Input/Output device 108 enables the central processor
15 unit 101 to communicate with or monitor other electronic systems, e.g., client devices displaying viewing windows, coupled to a communication network.

20 GENERAL METHOD AND SYSTEM FOR MULTI-USER VIDEO NAVIGATION THROUGH
A VIDEO SEQUENCE

As an overview, embodiments of the present invention provide a method and system for multi-user video navigation through a video sequence. The present invention further provides a method and system that overcomes the singular
25 functional control of video presentation for a video sequence by generating multiple viewing windows in a video sequence. Additionally, the present invention provides a method and system for functional control of video presentation over a video sequence that is scalable.

With reference now to Figure 2 and flow chart 700 of Figure 7, exemplary steps used by the various embodiments of the present invention are illustrated. Flow chart 700

includes processes of the present invention which, in one embodiment, are carried out by a processor under the control of computer-readable and computer executable instruction. The computer-readable and computer-executable instructions reside, for example, in data storage features such as computer usable volatile memory 102, computer usable non-volatile memory 103, and/or data storage device 104 of Figure 1. The computer-readable and computer executable instructions are used to control or operate in conjunction with, for example, central processing unit 101 of Figure 1.

With reference again to Figure 2, a diagram of an exemplary communication system 200 capable of capturing a video sequence and providing multiple and navigable windows of the video sequence to a plurality of client devices that are remotely located as employed in conjunction with various embodiments of the present invention is shown. The communication system 200 allows one or more client devices to display portions of a video sequence that is live-captured or stored.

The system 200 contains a single and stationary video camera 210, a video capturing module 220, a video compression module 270 and a server 230. The video camera captures a high-resolution, full-sized video sequence having

dimension in the x and y axis, in accordance with one embodiment. The original video sequence contains a sequence of individual picture frames, wherein each frame contains a still image. When the picture frames are shown together sequentially, the video sequence of picture frames is presented.

In the case of live-captured video, the capturing module 220 samples the signal from the video camera and produces pixel information for each frame of the video sequence forming a captured video stream. Essentially, the video stream produced by the capture module 220 is an uncompressed version of the original video sequence. On the other hand, the compression module 270 separately encodes and produces compressed bit streams of the original video sequence, forming a compressed video stream.

For stored video, a decoder or digitizer module (not shown) replaces the capture module 220 to obtain the digital domain pixel information. The resulting video stream produced by the decoder or digitizer module is an uncompressed version of the original video sequence.

The captured video sequence and the compressed video stream must be synchronized appropriately. In the least complicated situation, the compression module 270 allows the input signal of uncompressed video frames from the video camera 210 to be passed through separately for processing. Otherwise, in one embodiment, if the

compression module 270 does not allow such a pass-through, then a separate hardware video capture unit (e.g., the capture module 220) must be used and synchronized with the compressed bit stream.

5

The server 230 contains a pixel picture buffer 232, a coded picture buffer 236, and a video cropping engine 234 that are coupled together. The server 230 is responsible for receiving the control requests from client devices and generating the compressed video bit streams for the portions of the original video sequence that are transmitted to the requesting client devices. For every reconstructed portion of the video sequence, the transmitted compressed bit stream represents a smaller sized video sequence, which is a windowed portion of the full-sized video sequence.

As mentioned above, in the following discussion, the term "server" is in various embodiments intended to encompass a device functionally resembling a computer (e.g., having computation ability, memory, and/or connectivity capability). A typical server according to the definition as used in the present application may include, but is not limited to, any computer (e.g., mainframe, corporate server, personal computer (PC), laptop, and the like).

Contained within the server 230, the pixel picture buffer 232 receives the output from the capture module 220,

the uncompressed version of the original video sequence. The coded picture buffer 236 receives the output from the compression module 270, the compressed version of the original video sequence.

5

With reference now to Figure 7, flow chart 700 illustrates steps performed in accordance with one embodiment of the present invention are shown. Although flow chart 700 describes steps that are associated with one window in a video sequence, the steps in flow chart 700 are applicable in supporting multiple windows in a video sequence that are independently generated and navigated.

10

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209

corresponding to the requested cropped windows. The present embodiment then compresses the generated bit streams in the same compression format used in the compression module 270 and transmits the compressed bit streams to respective requesting client devices.

In step 720, the present embodiment generates a first window video sequence that corresponds to the window at the first location. In step 720, the present embodiment, through the video cropping engine 234, crops the video sequence to generate the first window video sequence. Cropping of the video sequence into the requested portions is accomplished by extracting and combining respective portions of the uncompressed and compressed versions of the original video sequence. A second compression operation is performed on the uncompressed portion to code the uncompressed portion to the compression format used by the compression module 270.

Since each frame of the transmitted windowed video sequence is a partial scene or a portion of the original video sequence, the video cropping engine 234 can parse the compressed bit stream coming from the compression module 270 and extract those parts that can be reused. This exploits the already compressed bit stream that is calculated by the original compression operation at the compression module 270. In one embodiment, selection of the appropriate blocks of information from the original compressed bit stream is accomplished in real time in

order to synchronize with the pixel-domain frames
processed at the capture module 220.

Thereafter, the present embodiment parses information
5 or data from the uncompressed version of the video
sequence to complete the first window video sequence.
Data from the uncompressed version is coded in an intra
mode in order to comply with the compression format used
by the compression module 270, in accordance with one
10 embodiment.

In a multi-user environment, the original video
sequence can be independently cropped into a plurality of
cropped portions. Each of the cropped portions correspond
15 to individually generated cropped windows and in one
embodiment contemporaneously transmitted to their
corresponding client devices, in step 730.

The term "contemporaneously" in this Application is
20 defined as occurring coincidentally, jointly, concurrently,
almost simultaneously, or the like, such that each of the
corresponding client devices are receiving and viewing
their respective cropped window video sequences during the
same period of time.

25 In step 740, the present embodiment navigates the
window from the first location to a second location within
the dimension of the video sequence. The navigation is in
response to a navigation request 250 communicated from the

remotely located client device. The present embodiment provides a means for users of client devices to individually navigate and view different portions of the video sequence by controlling the navigation offset and size of their respective windows.

In step 750, the present embodiment generates a second window video sequence that corresponds to the window at the second location. In the multi-user environment, the present embodiment in step 750 crops the video sequence into the windowed portions by extracting and combining respective portions of the uncompressed and compressed versions of the original video sequence.

Again, a second compression operation is performed on the uncompressed portion to code the uncompressed portion into the compression format used by the compression module 270 in one embodiment. As discussed previously, the video cropping engine 234 can parse the compressed bit stream coming from the compression module 270 and extract those parts that can be reused. This exploits the already compressed bit stream that is calculated by the original compression operation at the compression module 270.

In addition, selections of the first window video sequence that are in compliance with the compression algorithm for the requested window size and location are also reused. Specifically, the present embodiment extracts prediction information from the first window video sequence

that are based on blocks contained within the window at the second location. Reference blocks that comply with the compression format used in module 270 for the window in the second location can be reused by implementing an adjusted motion vector. The adjusted motion vector is the vector sum of the navigation step vector and the old motion vector, as will be further discussed.

Thereafter, information from the uncompressed version of the video sequence is intra coded in the compression format used by the compression module 270 and substitutes the aforementioned information to complete the second window video sequence. In step 760, the present embodiment transmits the second window video sequence to the requesting client device. The client device is remotely located from the server 230.

One embodiment of the present invention is scalable to multiple users and provides each user with individual navigational functionality and control through efficient video compression and processing algorithms implemented in hardware and/or software. In this manner, individual navigational control over each of the cropped windows determines which part of the original video sequence is displayed to each of the requesting users at their respective client devices.

In another embodiment of the present invention, a single hardware or software compression of the high

resolution sequence is separately performed. As such, a compressed version of the single high-resolution video sequence is generated. In this way, a compressed version and an uncompressed version of the high-resolution video sequence is created. The resulting compressed bit stream and the captured video frames are combined to generate a new compressed bit stream for each portion of the video sequence that is transmitted to client devices. Each of the transmitted bit streams contains a requested window of the video sequence as requested by the user. This embodiment achieves improved computational performance by exploiting the coded information as calculated by the original compression operation.

If a hardware compression unit is used, only one viewpoint can be compressed per hardware compression unit. However, if a software compression is used, the supported number of viewpoints will be determined by the computing power of the electronic device, e.g., device 100. In essence, one software encode operation is needed for each window.

With reference now to the flow chart 500 of Figure 5 and Figure 3, steps for generating a bit stream that corresponds to a windowed portion of an original video sequence are shown, in accordance with one embodiment of the present invention. The window described in conjunction with flow chart 500 is stationary.

Figure 3 illustrates a block diagram of an exemplary frame 300 at time "t" in the original video sequence that shows a single cropped window 310, in accordance with the present embodiment. In flow chart 500, the cropped window 310 is not moving between frames (e.g., time "t" and "t+1") of the original video sequence. The frame 300, and correspondingly the original video sequence, has dimension in the x and y axis. Also, the cropped window 310 has dimension within the frame 300 as shown by the lengths l_x 350 and l_y 340 that defines a size for the cropped window 310. A reference point 380 defines the x and y coordinates for locating the relative position of the cropped window 310 within the dimensions of the frame 300.

Those well versed in the art understand that the frame 300 is broken down into unit blocks that are individually treated and compressed depending on the compression module 270 used. The unit blocks are divided into equally sized and non-overlapping blocks in the frame 300. For example, in accordance with one embodiment, the Moving Picture Expert Group (MPEG) video compression standard is used in the compression module 270. The unit block size for an MPEG video encoder is sixteen by sixteen (16x16) pixels. These unit blocks correspond to similarly positioned blocks within frames of the compressed uncompressed versions of the original version sequence.

Embodiments of the present invention incorporate a method of compression known to those in the art as block

based motion compensation or block matching. During block matching, on an individual basis, each target block in the current frame (e.g., frame 300) is compared with a past frame (e.g., the frame at "t-1") in order to find a

5 matching block or blocked area that is used as a replica. For purposes of this Application, the matching block is also referred to as a "reference block." The past frame has already been transmitted to the client device. When the current frame is reconstructed by the receiver at the
10 client device, among other processes, this matching block is substituted for the target block in the current frame.

The search can be conducted throughout the entire dimension of the past frame; however, the search is usually
15 restricted to a smaller search area that is centered around the position of the target block in the current frame (e.g., frame 300 at time "t"). This search area is called the motion vector search range. The motion vector search range places an upper limit on how far objects can move
20 between frames to provide effective coding.

Once the best substitute, or matching block is found for the target block, a motion vector is calculated. The motion vector describes the location of the matching block
25 from the past frame in relation to the position of the target block in the current frame. For coding efficiency, the difference between the target block and the matching block is then encoded along with the motion vector and transmitted to the client device.

The cropped window 310 is comprised of the same unit blocks that correspond to the blocks in the frame 300. The blocks in the cropped window 310 can be further classified as boundary blocks 320 and non-boundary blocks 330. The boundary blocks 320 are the blocks that are located on the boundary of the cropped window 310. As such, the boundary blocks are located on the edges or borders of the cropped window 310. In addition, the boundary blocks 320 have motion vectors that point to reference blocks that may lie, either partially or entirely, outside the cropped window 310. Non-boundary blocks 330 lie within a region that is surrounded by the boundary blocks 320 in the cropped window 310. Further, non-boundary blocks 330 have motion vectors that definitely point to reference blocks that lie entirely within the cropped window 310.

Boundary blocks 320 can be further defined as useable and non-useable boundary blocks. For example, in Figure 3, block 370 lies on the edge of the cropped window 310 and has a motion vector 372 that points to a reference block (not shown) that lies outside of the cropped window 310. As such, block 370 is defined as a boundary block, and is further defined as a "non-useable" boundary block. Also, boundary block 380 has a motion vector 382 that points to a reference block (not shown) that lies entirely within the cropped window 310. As such, boundary block 380 is defined as a "useable" boundary block.

Figure 5 is a flow chart illustrating the method for generating the compressed bit stream associated with the cropped window 310 in Figure 3, in accordance with one embodiment of the present invention. Since the rendering of the generated compressed bit stream for the cropped window 310 must be done in real time, the information from the compression module (e.g., module 270 in Figure 2) is used as much as possible.

While it is straightforward to reuse non-boundary blocks 330 in the middle of the cropped scene, reusing some of the boundary blocks 320 may pose a problem. Some boundary blocks 320 of the cropped viewing window 310 may need to be further coded in an intra mode in one embodiment. These boundary blocks, that are recorded in intra mode, originally have prediction dependencies based on image data that lie outside of the cropped viewing window 310. However, image data in blocks that lie outside of the cropped viewing window 310 is not transmitted in the windowed video bit stream. As such, boundary blocks transmitted with dependencies on nontransmitted image data cannot be recreated at the client device. Independently encoding those boundary blocks in the intra mode allows those boundary blocks to be displayed at the client device without any loss of data.

In generating the windowed video bit stream for the portion of the original video sequence that corresponds to the cropped window 310, the present embodiment begins the

process outlined in flow chart 500 at the start of each
frame in the original video sequence, in step 510. In
Figure 5, the process outlined in flow chart 500 repeats
for every block within the cropped window 310, as is shown
in step 515.

In step 520, the present embodiment determines if the
block in question is a boundary block within the cropped
window 310. As discussed previously, a boundary block lies
on the edge of the cropped window and has prediction
dependencies based on image data that may lie outside of
the cropped viewing window 310.

If it is not a boundary block, the present embodiment
proceeds to step 545 and crops the block from the
compressed version generated by the compression module
(e.g., module 270). Since the bit stream generated by the
compression module is of the same compression format as the
video bit stream that is transmitted to the client device,
the present embodiment reuses or copies the part of the bit
stream from the compression module pertaining to the block
in question.

On the other hand, for boundary blocks where the
motion vector is pointing outside of the cropped window
310, the block is independently coded in one embodiment.
If necessary, pixel information for the boundary blocks can
be obtained from the capturing module 220. Thus, the
present embodiment determines if the boundary block, and

correspondingly, the entire frame, is intra coded, in step 525. An intra coded frame (I-frame) can be reconstructed without any reference to other frames. In other words, intra coded frames are independently coded.

5

If the target boundary block is intra coded, then the present embodiment proceeds to step 545 and crops the target boundary block from the compressed bit stream generated by the compression module (e.g., module 270).

10

The present embodiment proceeds to step 530, if the target boundary block is inter-coded. Inter-coded P-frames contain blocks that are reconstructed with data from blocks in a previous frame. A predicted frame (P-frame) is forward predicted, in that a previous I or P frame is used to reconstruct the current frame. Inter-coded B-frames are both forward and backward predicted from the last or next I or P frame. The B-frames require at least two frames in order to reconstruct the target blocks from the current frame.

20

In step 530, the present embodiment determines if the reference block pointed to by the motion vector lies within the cropped window. In other words, in step 530, the present embodiment determines if the boundary block is useable or non-useable.

25

If the boundary block is non-useable, then the present embodiment proceeds to step 535 and crops information

pertaining to the target boundary block from the uncompressed bit stream generated or passed by the capture module (e.g., module 220). Then the present embodiment encodes the information as an intra block in the

5 compression format used by the compression module (e.g., module 270). In this way, the video bit stream transmitted to the client device maintains the integrity of its compression format.

10 On the other hand, if the boundary block is useable, then the present embodiment proceeds to step 550 and adjusts the syntax for the compressed bit stream from the compression module (e.g., module 270) if necessary. For a boundary block, as long as its motion vector points to a
15 reference block that is entirely within the cropped window, that block information can also be reused.

However, certain syntax for the generated bit stream must be adjusted accordingly. For instance, when the
20 output stream is generated as an MPEG sequence, a boundary block may be a skipped block. However, the MPEG sequence does not allow the first block of the window or frame to be a skipped block. While the block may be appropriately coded for the entire frame, when cropping to the window
25 310, the block at that time may not comply with the MPEG standard. In that case, the coding syntax for the skipped block must be adjusted to comply with the MPEG standard in step 550. Thereafter, the present embodiment copies the adjusted block in step 555.

In step 560, before sending the video bit stream to the client device, from steps 540, 545, and 555, the present embodiment quantizes the generated video bit stream for the cropped window (e.g., 310) by a rate control module (not shown). The bit rate of the cropped video bit stream should be adjusted and maintained. The rate control module adjusts the transmitted bit rate to reflect the capabilities of the networking environment and the client device. Since the windowed video has smaller resolution than that of the original video sequence, it should be coded in a smaller bit rate. The rate control is done on the compressed data obtained from the compression module. Adjusting the bit rate is accomplished before the generated video bit stream for the cropped window is written to the output in step 545.

Figure 6 is a flow chart that, in combination with Figure 4, illustrates a method for generating the compressed bit stream at the moment an associated cropped window (e.g., window 310) is changing location within the original video sequence, in accordance with one embodiment of the present invention. While the user is moving a cropped viewing window from a first location to a second location within the dimensions of the original video sequence, there will be reference shifting during navigation of the cropped viewing window. In essence, the relative reference must be adjusted when navigation occurs during a prediction frame in order to reuse the motion

information already generated by the compression module
(e.g., module 270).

Figure 4 is a block diagram illustrating the movement
5 of a block in a cropped window (e.g., window 310) from an
original video sequence from a first location in a frame at
time "t" to a second location in a frame at time "t+1." In
Figure 4, at frame "t", the block is denoted as the "old"
block 415. The new position of the same block in the
10 viewing window at frame "t+1" is represented as "new" block
425.

The movement of the block represents the same movement
of the cropped window as requested by the user, via a
15 navigation control request, at a client device. A
navigation step (s_x, s_y) 430 represents movement of the
block, and the associated cropped window in one embodiment.
The navigation step can be represented also as navigation
step vector 430.

20 The navigation step size, representing the horizontal
and vertical movement of the selected cropped scene, is a
multiple of the size of whatever the motion estimation
unit that is used in the compression module (e.g., module
25 270), in accordance with one embodiment of the present
invention. For example, if the compression module is an
MPEG video encoder, motion estimation is typically done in
unit blocks of size sixteen pixels by sixteen pixels

(16x16 pixels). As such, the navigation step size in the x and y coordinates is a multiple of 16 pixels.

In accordance with another embodiment, considering a viewing window shifting from time "t" to "t+1" as shown in Figure 4, in general, if the frame at time t+1 is an intra frame, the new block 425 is cropped directly from the compression module. The information for the corresponding blocks coded from the compression module are copied to the output video bit stream to the client, since they are coded independently.

On the other hand, if the frame is a predicted frame, in order to reuse block information, an associated motion vector must be adjusted since there is shifting between the reference frame and the current frame (t+1) due to the navigation step 430, in accordance with one embodiment of the present invention. For example, the motion vector associated with the old block 415 points to a reference block 460. The reference block is within the "old" motion vector search range 410 and can be properly used.

Two conditions exist for reusing the blocks within the frames at time "t" in one embodiment. The first condition is that the adjusted motion vector is within the new motion vector search range (e.g., search range 420). In other words, the adjusted motion vector points to a reference block such that is the adjusted motion vector is within the new motion vector search range (e.g., search range 420).

In one embodiment, the motion vector search range is within a set of [-16 pixels, +16 pixels] in both the x and y coordinates. The second condition is that any part of the reference block should not be outside of the cropped window at time "t."

In another embodiment, if the boundary of the cropped viewing window aligns with the boundary of the original scene of the original video sequence, the motion vector will always be within the range of the new motion vector search range. Therefore, information regarding those boundary blocks can always be reused. In accordance with another embodiment, navigations outside of the original scene are not possible.

Flow chart 600 of Figure 6 shows the processing flow for the first frame when the position of the cropped viewing window is changed from one location to another, in accordance with one embodiment of the present invention.

The user at a client device sends the location and size of the new viewing window in a navigation control request. The present embodiment computes the navigation step by obtaining the differences between the locations of the viewing window at time "t" and "t+ 1," in step 610.

In step 620 of flow chart 600, the present embodiment then waits until the next start of a picture frame, the frame at time "t+1," by looking into the coded picture buffer (e.g., buffer 236 in Figure 2) which is filled by

the compression module (e.g., module 270). The corresponding pixel data for the same picture should be contained in the pixel picture buffer (e.g., buffer 232) which is filled by the capturing module.

5

Thereafter, the present embodiment completes the following steps for each block as indicated in step 620. In condition step 620, the present embodiment determines if the block in question (block 425), and correspondingly, the entire frame, is intra coded. As discussed previously, an intra coded frame (I-frame) can be reconstructed without any reference to other frames.

10

15

For each compressed block in the picture, if it is independently coded, it can be used directly without any change. As such, if the block is intra coded, then the present embodiment proceeds to step 640 and crops the new block 425 from the compressed bit stream generated by the compression module (e.g., module 270).

20

The present embodiment proceeds to step 630, if the new block 425 is inter coded. As discussed previously, inter coded P-frames contain blocks that are reconstructed with data from blocks in a previous frame. Inter coded B-frames are both forward and backward predicted from the last or next I or P frame. The B-frames require at least two frames in order to reconstruct the target blocks from the current frame. In the present embodiment, a cropped window cannot be moved during the interpolated B-frames.

25

If it is predicted from the frame at time t , the present embodiment in flow chart 600 checks to see if the aforementioned conditions are satisfied in order to reuse the coded block in the frame at time " $t+1$." Otherwise, the corresponding pixel block is cropped from the pixel picture and encoded as an intra block.

In step 630, the present embodiment determines if the navigation step vector is greater than the motion vector search range.

The present embodiment in flow chart proceeds to step 650 if the navigation step vector is greater than the motion vector search range and crops information pertaining to the new block 425 from the uncompressed bit stream generated or passed by the capture module (e.g., module 220). Then, the present embodiment encodes the information as an intra block in the compression format used by the compression module (e.g., module 270). In this way, the video bit stream transmitted to the client device maintains the integrity of its compression format.

However, if the navigation step vector is not greater than the motion vector search range, then the present embodiment adjust the motion vector according the navigation step in step 635. For example, referring back to Figure 4, in order to reuse the information coded in a predicted frame at time " $t+1$," the old motion vector 440

must be adjusted to reflect the new position of the new block 425 in relation to the same reference block 460. In other words, the new motion vector 450 is adjusted to the new position of the new block 425. The new adjusted motion vector can be computed using vector analysis. In one embodiment, the adjusted motion vector can be computed as the sum of the old motion vector (d_x, d_y) 440 and the navigation step vector (s_x, s_y) 430.

In x and y coordinates, the adjusted motion vector (d'_x, d'_y) can be computed as in the x and y coordinates, as follows:

$$d'_x = d_x + s_x$$

$$d'_y = d_y + s_y$$

In accordance with one embodiment of the present invention, the adjusted new motion vector must be within the motion vector search range 420 of the new block 425. In condition step 645, the present embodiment determines if the adjusted motion vector points to a reference block that is within the motion vector search range for the new block 425.

In the case where the adjusted motion vector is within the motion vector search range, the present embodiment copies the adjusted block in step 670. In Figure 4, reference block 460 is within the motion vector search range 420 and can be used.

On the other hand, if the adjusted motion vector is not within the motion vector search range, then the present embodiment would proceed to step 650 as discussed previously. By way of example, if block 470 were to be used as a reference block, then block 470 would be outside of the motion vector search range 420 and could not be used. This is the case even though the block 470 was originally within the old motion vector search range 410.

In step 675, before sending the video bit stream to the client device, from steps 640, 655, and 670, the present embodiment quantizes the generated video bit stream for the cropped window by a rate control module (not shown). The bit rate of the cropped video bit stream should be adjusted and maintained. The rate control module adjusts the transmitted bit rate to reflect the capabilities of the networking environment and the client device. Since the windowed video has smaller resolution than that of the original video sequence, it should be coded in a smaller bit rate. The rate control is done on the compressed data obtained from the compression module. Adjusting the bit rate is accomplished before the generated video bit stream for the cropped window is written to the output in step 680.

While the methods of embodiments illustrated in flow charts 500, 600, and 700 show specific sequences and quantity of steps, the present invention is suitable to alternative embodiments. For example, not all the steps

provided for in the method are required for the present invention. Furthermore, additional steps can be added to the steps presented in the present embodiment. Likewise, the sequences of steps can be modified depending upon the application.

A method for providing navigation for cropped viewing windows in a video sequence that are transmitted contemporaneously, is thus described. While the present invention has been described in particular embodiments, it should be appreciated that the present invention should not be construed as limited by such embodiments, but rather construed according to the below claims.